2019 American Control Conference (ACC)
Philadelphia, PA, USA, July 10-12, 2019

The Impacts of Behavioral Probability Weighting on Security Investments in Interdependent Systems

Mustafa Abdallah, Parinaz Naghizadeh, Ashish R Hota, Timothy Cason,
Saurabh Bagchi, and Shreyas Sundaram.

Abstract— **We consider a system consisting of multiple interdependent assets, and a set of defenders, each responsible for securing a subset of the assets against an attacker. The interdependencies between assets are captured by an attack graph, where an edge from one asset to another indicates that if the former asset is compromised, an attack can be launched on the latter asset. Each edge has an associated probability of successful attack, which can be reduced via security investments by the defender responsible for that edge. While prior work has studied the security investments in such scenarios, in this work we consider what happens when the defenders exhibit characteristics of boundedly-rational human decision-making that have been identified by behavioral economics. In particular, humans have been shown to perceive probabilities in a nonlinear manner, typically overweighting low probabilities and underweighting high probabilities. We show that such nonlinear probability weighting can affect the security investments in interdependent systems, and suboptimal investments can arise under such weighting in certain network topologies. We also show that the presence of a defender who exhibits behavioral probability weighting can be beneficial for the other defenders in the network, in terms of making their assets more secure.**

I. Introduction

Large-scale cyber-physical systems (CPS) consist of multiple interdependent subsystems managed by different stakeholders (or operators). Such systems are increasingly under threat from sophisticated adversaries who seek to compromise various assets in the system. As a result, there has been a significant amount of research dedicated to understanding how to improve the security of such systems [1], [2]. In particular, game theory has played a key role in reasoning about security problems, due to its ability to systematically capture the potentially conflicting goals of the various actors (e.g., defenders and attackers) in the system [3]–[5].

In the context of large-scale interdependent systems, adversaries can often use stepping-stone attacks to exploit vulnerabilities within the network in order to compromise a particular target. Such threats can be captured via the notion of *attack graphs* that represent all possible paths that attackers may have to reach their targets within the CPS [6]. The defenders in such systems are each responsible for defending some subset of the assets [3], [7], and usually have limited resources (i.e., budget) that they can use to mitigate vulnerabilities in the network. These settings have been explored under various assumptions on the strategies available to the defenders and attackers [7]–[9].

A common thread in the existing literature is that the defenders are assumed to behave according to classical models of fully rational decision-making, taking actions to minimize their expected loss. However, a large body of work in behavioral economics and psychology has shown that humans consistently deviate from such classical models of decision-making. For example, *prospect theory* (introduced by Kahneman and Tversky in their seminal paper [10]) showed that humans perceive gains, losses and probabilities in a skewed (nonlinear) manner, typically overweighting low probabilities and underweighting high probabilities. While a large literature on prospect theory exists in economics and psychology, relatively little research has investigated the effect of such behavioral decision-making on CPS security and robustness (exceptions include [11]–[13]).

In this paper, we introduce prospect theory into a game-theoretic framework involving attack graph models of large-scale interdependent systems with multiple defenders. Specifically, we consider the scenario where each (human) defender misperceives the probabilities of successful attack in the attack graph. We characterize the impacts of such misperceptions on the security investments made by each defender. In contrast to [11], we consider a more general case in which each defender is responsible for a subnetwork (i.e., set of assets) rather than just a single node, and where the attacker exploits paths through the network to reach certain target nodes.

We introduce a setting for behavioral decision-making, described in further detail in the next two sections. We first establish the convexity of the objective function of each defender, and we use this to prove the existence of a pure strategy Nash equilibrium (PNE). We then characterize the impacts of probability weighting on the investment decisions made by the defenders; in particular, we show that nonlinear perceptions of probability can induce defenders to invest in a manner that increases the vulnerability of their assets to attack. We characterize classes of graphs where such effects arise. Furthermore, we illustrate the impacts of having a mix of defenders (with heterogeneous levels of probability weighting bias) in the system, and show that the presence of defenders with skewed perceptions of probability can in fact *benefit* the non-behavioral defenders in the system.

This research was supported by grant CNS-1718637 from the National Science Foundation. Mustafa Abdallah, Parinaz Naghizadeh, Saurabh Bagchi, and Shreyas Sundaram are with the School of Electrical and Computer Engineering at Purdue University. Email: {abdalla0,parinaz,sbagchi,sundara2}@purdue.edu.
Ashish R. Hota is with the Department of Electrical Engineering, Indian Institute of Technology (IIT), Kharagpur, India. Email: ahota@ee.iitkgp.ac.in. Timothy Cason is with the Krannert School of Management at Purdue University. Email: cason@purdue.edu.

978-1-5386-7926-5/$31.00 ©2019 AACC

II. The Security Game Framework

In this section, we describe our general security game framework, including the attack graph and the characteristics of defenders and attackers.

A. Attack Graph

We represent the assets in a CPS as nodes of a directed graph $G = (V, \mathcal{E})$ where each node $v_i \in V$ represents an asset. A directed edge $(v_i, v_j) \in \mathcal{E}$ means that if node v_i is successfully attacked, it can be used to launch an attack on node v_j. We assume that the success of attacks across different edges in the network are captured by independent random variables. Each edge $(v_i, v_j) \in \mathcal{E}$ has an associated weight $p^0_{i,j} \in (0, 1]$, denoting the probability of successfully attacking asset v_j starting at v_i (in the absence of any security investments). The graph contains a designated source node v_s, which is used by the attacker to begin her attack on the network. For a general asset $v_t \in V$, we define $\mathcal{P}_t$ to be the set of directed paths from the source v_s to v_t on the graph, where a path $P \in \mathcal{P}_t$ is a collection of edges $\{(v_s, v_1), (v_1, v_2), ..., (v_k, v_t)\}$. Therefore, in the absence of any security investments, the probability that v_t is compromised due to an attacker exploiting a given path $P \in \mathcal{P}_t$ is $\prod_{(v_m, v_n) \in P} p^0_{m,n}$, by our aforementioned independence assumption. The attacker can choose any path from the multiple attack paths in $\mathcal{P}_t$ to attack v_t.

B. Strategic Defenders

Let $\mathcal{D}$ be the set of all defenders of the network. Each defender $D_k \in \mathcal{D}$ is responsible for defending a set $V_k \subseteq V \setminus \{v_s\}$ of assets. For each compromised asset $v_m \in V_k$, the defender D_k will incur a financial loss $L_m \in \mathbb{R}_{\geq 0}$. To reduce the attack success probabilities on edges interconnecting assets inside the network, a defender can allocate security resources on these edges, subject to the constraints described below.

Let $\mathcal{E}_k \subseteq \mathcal{E}$ be the subset of edges that defender D_k can allocate security resources on, with $n_k = |\mathcal{E}_k|$. We assume that each defender D_k has a security budget $B_k \in \mathbb{R}_{\geq 0}$. Thus, we define the defense strategy space of each defender $D_k \in \mathcal{D}$ by

$$X_k \triangleq \{x^k_{i,j} \in \mathbb{R}_{\geq 0}, (v_i, v_j) \in \mathcal{E}_k : \sum_{(v_i, v_j) \in \mathcal{E}_k} x^k_{i,j} \leq B_k\}. \quad (1)$$

In words, the defense strategy space for defender D_k consists of all nonnegative investments on edges under her control, with the sum of all investments not exceeding the budget B_k. We denote any particular vector of investments by defender D_k by $x_k \in X_k$.

Let $\mathbf{x} = \left[x_1, x_2, \ldots, x_{|\mathcal{D}|}\right]$ be a joint defense strategy of all defenders, where $x_k \in X_k$ for every defender D_k. Under a joint defense strategy $\mathbf{x}$, the total investment on edge (v_i, v_j) is $x_{i,j} := \{\sum_{D_k \in \mathcal{D}} x^k_{i,j} : (v_i, v_j) \in \mathcal{E}_k\}$. Let $p_{i,j} : \mathbb{R}_{\geq 0} \to [0, 1]$ be a function mapping the total investment $x_{i,j}$ to an attack success probability, and with $p_{i,j}(0) = p^0_{i,j}$.

The goal of each defender D_k is to choose her investment vector x_k in order to best protect her assets from being attacked. In this paper, we consider the scenario where each defender minimizes the highest probability path to each of her assets; this captures settings where the specific path taken by the attacker is not known to the defender a priori, and thus the defender seeks to make the most vulnerable path to each of her assets as secure as possible. Mathematically, this is captured via the cost function

$$C_k(\mathbf{x}) = \sum_{v_m \in V_k} L_m \Big(\max_{P \in P_m} \prod_{(v_i, v_j) \in P} p_{i,j}(x_{i,j}) \Big) \quad (2)$$

subject to $x_k \in X_k$. Note that $C_k(\mathbf{x})$ is a function of the investments of all defenders, and thus we denote the cost by $C_k(x_k, \mathbf{x}_{-k})$ where $\mathbf{x}_{-k}$ is the vector of investments by defenders other than D_k. Each defender chooses her investment vector $x_k \in X_k$ to minimize the cost $C_k(x_k, \mathbf{x}_{-k})$, given the investments $\mathbf{x}_{-k}$ by the other defenders.

The recent work [9] studies the above security game setting assuming rational defenders, and provides a method to calculate the optimal investments by the defenders with respect to the cost function (2). However, as mentioned in the introduction, humans have been shown to systematically misperceive probabilities, which can impact the decisions that they make in the presence of risk. In the next section, we will review certain classes of probability weighting functions that capture this phenomenon, and then subsequently introduce such functions into the above security game formulation. We will then characterize how such behavioral decision-making affects the security investments made by the defenders in this game.

III. Nonlinear Probability Weighting and the Behavioral Security Game

A. Nonlinear Probability Weighting

The behavioral economics and psychology literature has shown that humans consistently misperceive probabilities by overweighting low probabilities and underweighting high probabilities [10], [14]. More specifically, humans perceive a "true" probability $p \in [0, 1]$ as $w(p) \in [0, 1]$, where $w(\cdot)$ is a probability weighting function. A commonly studied probability weighting function was proposed by Prelec in [14], and is given by

$$w(p) = \exp\Big[-(-\log(p))^\alpha\Big], \quad p \in [0, 1], \quad (3)$$

where $\alpha \in (0, 1]$ is a parameter that controls the extent of overweighting and underweighting. When $\alpha = 1$, we have $w(p) = p$ for all $p \in [0, 1]$, which corresponds to the situation where probabilities are perceived correctly. Smaller values of α lead to a greater amount of overweighting and underweighting, as illustrated in Fig. 1. Next, we incorporate this probability weighting function into the security game defined in the last section, and define the Behavioral Security Game that is the focus of this paper.

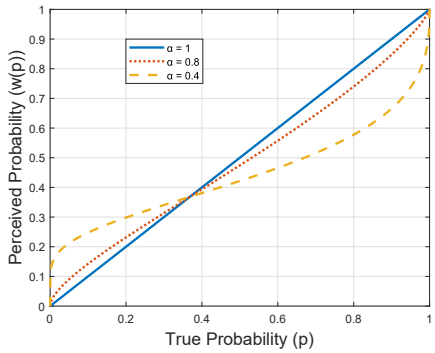

Fig. 1: Prelec Probability weighting function (3) which transforms true probabilities p into perceived probabilities $w(p)$. The parameter α controls the extent of overweighting and underweighting.

B. The Behavioral Security Game

Recall that each defender seeks to protect a set of assets, and the probability of each asset being successfully attacked is determined by the corresponding probabilities on the edges that constitute the paths from the source node to that asset. This motivates a broad class of games that incorporate probability weighting, as defined below.

Definition 1: We define a *Behavioral Security Game* as a game between different defenders in an interdependent network, where each defender misperceives the attack probability on each edge according to the probability weighting function defined in (3). Specifically, the perceived attack probability by a defender D_k on an edge (v_i, v_j) is given by

$$w_k(p_{i,j}(x_{i,j})) = \exp\Big[-(-\log(p_{i,j}(x_{i,j})))^{\alpha_k}\Big],$$

where $p_{i,j}(x_{i,j}) \in [0,1]$ and $\alpha_k \in (0,1]$.

Remark 1: The subscript k in α_k and $w_k(\cdot)$ allows each player in the Behavioral Security Game to have a different level of misperception.

Incorporating this into the cost function (2), each defender D_k seeks to minimize her *perceived* loss via the cost function

$$C_k(x_k, \mathbf{x}_{-k}) = \sum_{v_m \in V_k} L_m \left(\max_{P \in \mathcal{P}_m} \prod_{(v_i,v_j)\in P} w_k\left(p_{i,j}(x_{i,j})\right) \right). \tag{4}$$

IV. Existence of a Pure Nash Equilibrium

We start by establishing the existence of a pure strategy Nash Equilibrium (PNE) for the class of behavioral games defined in the previous section. A profile of security investments by the defenders is said to be a PNE if no defender can decrease her cost by unilaterally changing her security investment. We first prove the convexity of the defenders' cost function, and subsequently show the existence of a PNE for the Behavioral Security Game. Throughout, let the function $p_{i,j}(x_{i,j})$ represent the probability of successful attack on an edge when the total investment on that edge is $x_{i,j}$.

Assumption 1: For every edge (v_i, v_j), the probability of successful attack $p_{i,j}(x_{i,j})$ is a twice continuously differentiable, log-convex[1], and decreasing function.

Lemma 1: For every edge (v_i, v_j), the perceived probability $w(p_{i,j}(x_{i,j}))$ is convex $\forall x_{i,j} \in \mathbb{R}_{\geq 0}$ under Assumption 1.

Proof: First note that since $0 \leq p_{i,j}(x_{i,j}) \leq 1$, we have $0 \leq -\log(p_{i,j}(x_{i,j})) \leq \infty$ for all $x_{i,j} \in \mathbb{R}_{\geq 0}$. We drop the subscript i,j in the following analysis for better readability. Substituting into the probability weighting function defined in (3), we have

$$\begin{aligned} w(p(x)) &= \exp\Big[-(-\log(p(x)))^{\alpha}\Big] \\ &= (g \circ h)(x), \end{aligned}$$

where $g(x) = \exp(-x)$ and $h(x) = (-\log(p(x)))^{\alpha}$. Next, we prove that $h(x)$ is concave:

$$\begin{aligned} h'(x) &= -\alpha(-\log(p(x)))^{\alpha-1}\frac{p'(x)}{p(x)} \\ h''(x) &= \alpha(\alpha-1)(-\log(p(x)))^{\alpha-2}\frac{(p'(x))^2}{(p(x))^2} \\ &\quad + \alpha(-\log(p(x)))^{\alpha-1}\left[\frac{(p'(x))^2 - p(x)p''(x)}{(p(x))^2}\right]. \end{aligned}$$

Since $0 < \alpha \leq 1$, the first term on the R.H.S. of $h''(x)$ is non-positive. Also, since $p(x)$ is twice-differentiable and log-convex with a convex feasible defense strategy domain $\mathbb{R}_{\geq 0}$, $(p'(x))^2 \leq p(x)p''(x)$ [17], which ensures that the second term is also non-positive. Therefore, $h(x)$ is concave.

Finally, since $g(x)$ is convex and non-increasing while $h(x)$ is concave, $w(p(x))$ is convex. ■

Using the above result, we now prove that the defender cost function in the Behavioral Security Game is convex.

Lemma 2: Under Assumption 1, the cost function (4) of the Behavioral Security Game is convex for any $0 < \alpha_k \leq 1$.

Proof: Beginning with the cost function defined in (4), $w_k(p_{i,j}(x_{i,j}))$ is convex as shown in Lemma 1. Since the product of convex functions is convex if all of them are non-increasing (or non-decreasing) and positive on an interval [17] and $w_k(p_{i,j}(x_{i,j}))$ is monotone (composition of two monotonic functions), $\prod_{(v_i,v_j)\in P} w_k(p_{i,j}(x_{i,j}))$ is convex. Moreover, the maximum of a set of convex functions is also convex [18]. Finally, since the cost function $C_k(\mathbf{x})$ is a linear combination of convex functions, $C_k(\mathbf{x})$ is convex. ■

This brings us to the following result, establishing the existence of a PNE in the Behavioral Security Game.

Theorem 1: The Behavioral Security Game possesses a pure strategy Nash equilibrium (PNE) when $\alpha_k \in (0,1]$ for each defender D_k, and under Assumption 1.

Proof: The feasible defense strategy space X_k in (1) is nonempty, compact and convex for each defender D_k. Furthermore, as shown in Lemma 2, the cost function of

[1]This is a common assumption in the literature. In particular, [15] shows that log-convexity of the attack probability functions is a necessary and sufficient condition for the optimal security investment result of the seminal paper [16] to hold.

each defender is convex under the given assumptions. As a result, the Behavioral Security Game is an instance of *concave games*, which always have a PNE [19]. ■

In the following sections, we will provide further insights into the investments in the PNE of such games. In order to maintain analytical tractability, we will assume henceforth that the probabilities of successful attack on each edge are given by

$$p_{i,j}(x_{i,j}) = p^0_{i,j} \exp\Big(- \sum_{D_k \in \mathcal{D}} x^k_{i,j} \Big). \tag{5}$$

In other words, the probability of successful attack on an edge decreases exponentially with the sum of the investments by all players on that edge. Such probability functions fall within the class commonly considered in security economics (e.g., [16]), and satisfy the conditions in Lemma 1.

Consequently, the attack success probability of any given path P from the source to a target v_t is given by

$$\begin{aligned} Q(\mathbf{x}) &= \prod_{(v_m,v_n)\in P} p_{m,n}(x_{m,n}) \\ &= \Big(\prod_{(v_m,v_n)\in P} p^0_{m,n} \Big) \exp\Big(- \sum_{(v_m,v_n)\in P} \sum_{D_k \in \mathcal{D}} x^k_{m,n} \Big). \end{aligned} \tag{6}$$

Thus, the probability of successful attack on a given path decreases exponentially with the sum of the investments on all edges on that path by all players.

V. Properties of the Optimal Investment Decisions

In this section, we characterize properties of the optimal investment decisions under behavioral (i.e., $\alpha < 1$) and non-behavioral (i.e., $\alpha = 1$) decision-making. This setting will help in understanding the actions (i.e., best responses) of each behavioral defender in the Behavioral Security Games and their decisions under the PNE (note that a PNE always exists as proven in Section IV). We will subsequently return to the multiple defender game in the next section. In this section, we will refer to the defender as D (i.e., we will drop the index k throughout the section). We will identify how probability weighting affects the investments a defender makes to secure her assets against attacks.

A. Locations of Optimal Investments for Behavioral and Non-Behavioral Players

We first characterize the optimal investments by a non-behavioral player who is protecting a single asset, and subsequently compare that to the investments made by a behavioral player. In the following result, we use the notion of a *min-cut* in the graph. Specifically, given two nodes s and t in the graph, an edge-cut is a set of edges $\mathcal{E}_c \subset \mathcal{E}$ such that removing $\mathcal{E}_c$ from the graph also removes all paths from s to t. A min-cut is an edge-cut of smallest cardinality over all possible edge-cuts [20].

Theorem 2: Consider an attack graph $G = (V, \mathcal{E})$ where the initial attack success probabilities on all edges are equal to 1 (i.e., $p^0_{i,j} = 1,\ \forall (v_i, v_j) \in \mathcal{E}$). Let the attack success probability under security investments be given by $p_{i,j}(x_{i,j}) = e^{-x_{i,j}}$, where $x_{i,j} \in \mathbb{R}_{\geq 0}$ is the investment on edge (v_i, v_j). Suppose there is a single target asset v_t (i.e., all other assets have loss 0). Let $E_c \subseteq \mathcal{E}$ be a minimum edge cut set between the source node v_s and the target v_t. Then, it is optimal for a non-behavioral defender D to distribute all her investments equally only on the edge set E_c in order to minimize (2).

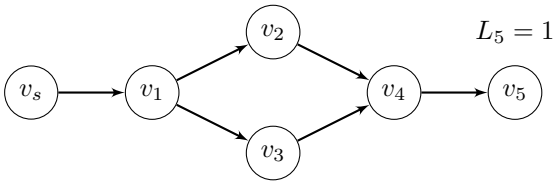

Fig. 2: An attack graph where a behavioral player makes suboptimal investment decisions.

Proof: Let E_c be a minimum edge cut between the source entry node v_s and the target node v_t. Also, let $N = |E_c|$ represent the number of edges in the minimum cut set E_c. Let B be the defender's budget.

Consider any optimal investment of that budget. Recall from (6) that the probability of a successful attack of the target along a certain path P is a decreasing function of the sum of the investments on the edges on that path. Using Menger's theorem [20], there are N edge-disjoint paths between v_s and v_t in G. At least one of those paths has total investment at most $\frac{B}{N}$. Therefore, the path with highest probability of attack from v_s to v_t has total investment at most $\frac{B}{N}$.

Now consider investing $\frac{B}{N}$ on each edge in the edge cut. Since every path from v_s to v_t goes through at least one edge in E_c, every path has at least $\frac{B}{N}$ in total investment. Thus, it is optimal to only invest on edges in E_c.

Finally, consider investing non-equally on edges in E_c where an edge $e_{i,j} \in E_c$ has investment $x_{i,j} < \frac{B}{N}$. Under this investment and since there are N edge-disjoint paths from v_s to v_t in G, $\exists$ a path P from v_s to v_t that has total investment less than $\frac{B}{N}$. Thus, the path with highest probability of attack P has a probability of attack larger than $\exp\big(-\frac{B}{N}\big)$ (which arises when investing $\frac{B}{N}$ equally on each edge in E_c). Therefore, the cost function in (2) is higher with this non-equal investment. Thus, the optimal investment on E_c must contain $\frac{B}{N}$ investment on each edge in E_c. ■

The above result of Theorem 2 no longer holds when we consider the investments by a behavioral player (i.e., with $\alpha < 1$), as illustrated by the following example.

Example 1: Consider the attack graph shown in Fig. 2, with a single defender D and a single target asset v_5 (with a loss of $L_5 = 1$ if successfully attacked). Let the defender's budget be B, and let the probability of successful attack on each edge (v_i, v_j) be given by $p_{i,j}(x_{i,j}) = e^{-x_{i,j}}$, where $x_{i,j}$ is the investment on that edge.

This graph has two possible min-cuts, both of size 1: the edge (v_s, v_1), and the edge (v_4, v_5). Thus, by Theorem 2,

it is optimal for a non-behavioral player to put all of her budget on either one of these edges.

Now consider a behavioral player with $\alpha < 1$. With the above expression for $p_{i,j}(x_{i,j})$ and using the Prelec function (3), we have $w(p_{i,j}(x_{i,j})) = e^{-x_{i,j}^{\alpha}}$. Thus, the cost function (4) for the Behavioral Security Game is given by

$$C(\mathbf{x}) = \max\Big(e^{-x_{s,1}^{\alpha}-x_{1,2}^{\alpha}-x_{2,4}^{\alpha}-x_{4,5}^{\alpha}}, \\ e^{-x_{s,1}^{\alpha}-x_{1,3}^{\alpha}-x_{3,4}^{\alpha}-x_{4,5}^{\alpha}}\Big),$$

corresponding to the two paths from the source v_s to the target v_t. One can verify (using the KKT conditions) that the optimal investments are given by

$$x_{1,2} = x_{2,4} = x_{1,3} = x_{3,4} = 2^{\frac{1}{\alpha-1}} x_{s,1} \ , \\ x_{4,5} = x_{s,1} = \frac{B - 4x_{1,2}}{2} = \frac{B}{2 + 4(2^{\frac{1}{\alpha-1}})} \ .$$

Thus, for the cost function (2), the optimal investments (corresponding to the non-behavioral player) yield a cost of e^{-B}, whereas the investments of the behavioral player yield a cost of $e^{-2^{\frac{\alpha}{\alpha-1}}} e^{-\frac{B}{1+2^{\frac{\alpha}{\alpha-1}}}}$, which is larger than that of the non-behavioral player for any $B > 2$. □

The above example illustrates a key phenomenon: as the defender's perception of probabilities becomes increasingly skewed (captured by α becoming smaller), she shifts more of her investments from the min-cut edges to the edges on the parallel paths between v_1 and v_4. This is in contrast to the optimal investments (made by the non-behavioral player) which lie entirely on the min-cut edges. Indeed, by taking the limit as $\alpha \uparrow 1$, we have

$$x_{i,j} = \lim_{\alpha \uparrow 1} \ 2^{\frac{1}{\alpha-1}} \ x_{s,1} = 2^{-\infty} \ x_{s,1} = 0$$

for edges (v_i, v_j) on the two parallel portions of the graph.

We now use this insight to identify graphs where a behavioral player makes suboptimal security investments.

Proposition 1: Consider an attack graph G with a source v_s and a target v_t. Let E_c be a minimum edge cut between v_s and v_t, with size $|E_c| = N$. Suppose the graph contains another edge cut E_c' such that $E_c' \cap E_c = \emptyset$, $|E_c'| > |E_c|$, and for each edge in E_c', there is a path from v_s to v_t that goes through that edge but none of the other edges in E_c'. Let the probability of successful attack on each edge $(v_i, v_j) \in \mathcal{E}$ be given by $p_{i,j}(x_{i,j}) = e^{-x_{i,j}}$, where $x_{i,j}$ is the investment on that edge. Let B be the budget of the defender. Then, if $0 < \alpha < 1$, investing solely on the min-edge-cut set E_c is not optimal from the perspective of a behavioral player.

Proof: Denote $M = |E_c'| > |E_c| = N$. By Theorem 2, it is optimal to invest the entire budget only on edges in E_c in order to minimize the cost function (2). We will show that this investment is not optimal with respect to the behavioral player's cost function (4).

Starting with the optimal investments on the min edge cut E_c where each edge in E_c has nonzero investment (as given by Theorem 2), remove a small investment ϵ from each of those N edges, and add an investment of $\frac{N\epsilon}{M}$ to each of the edges in E_c'. We show that when ϵ is sufficiently small, this will lead to a net reduction in perceived probability of successful attack on each path from v_s to v_t.

Consider any arbitrary path P from v_s to v_t. Starting with the investments only on the minimum edge cut E_c, the perceived probability of successful attack on path P will be

$$f_1(\mathbf{x}) = \exp\Big(-\sum_{\substack{(v_i,v_j)\in E_c,\\ (v_i,v_j)\in P}} x_{i,j}^{\alpha}\Big).$$

After removing ϵ investment from each of the N edges in E_c, and adding an investment of $\frac{N\epsilon}{M}$ to each of the edges in E_c', the perceived probability on path P will be:

$$f_2(\mathbf{x}) = \exp\Big(-\sum_{\substack{(v_i,v_j)\in E_c',\\ (v_i,v_j)\in P}} \Big(\frac{N\epsilon}{M}\Big)^{\alpha} - \sum_{\substack{(v_i,v_j)\in E_c,\\ (v_i,v_j)\in P}} (x_{i,j}-\epsilon)^{\alpha}\Big).$$

The net reduction in perceived probability on path P will be positive if $f_2(\mathbf{x}) < f_1(\mathbf{x})$, i.e.,

$$\sum_{\substack{(v_i,v_j)\in E_c',\\ (v_i,v_j)\in P}} \Big(\frac{N\epsilon}{M}\Big)^{\alpha} + \sum_{\substack{(v_i,v_j)\in E_c,\\ (v_i,v_j)\in P}} (x_{i,j}-\epsilon)^{\alpha} > \sum_{\substack{(v_i,v_j)\in E_c,\\ (v_i,v_j)\in P}} x_{i,j}^{\alpha}. \tag{7}$$

If we define

$$f(\epsilon) = \sum_{\substack{(v_i,v_j)\in E_c',\\ (v_i,v_j)\in P}} \Big(\frac{N\epsilon}{M}\Big)^{\alpha} + \sum_{\substack{(v_i,v_j)\in E_c,\\ (v_i,v_j)\in P}} (x_{i,j}-\epsilon)^{\alpha},$$

we see that inequality (7) is equivalent to showing that $f(\epsilon) > f(0)$. We have

$$\frac{df}{d\epsilon} = \frac{\alpha N}{M} \sum_{\substack{(v_i,v_j)\in E_c',\\ (v_i,v_j)\in P}} \Big(\frac{N\epsilon}{M}\Big)^{\alpha-1} - \alpha \sum_{\substack{(v_i,v_j)\in E_c,\\ (v_i,v_j)\in P}} (x_{i,j}-\epsilon)^{\alpha-1}.$$

Note that $\lim_{\epsilon\downarrow 0} \frac{df}{d\epsilon} = \infty$ which shows that $f(\epsilon)$ is increasing in ϵ for sufficiently small ϵ. Therefore, $f_2(\mathbf{x}) < f_1(\mathbf{x})$ for sufficiently small ϵ. Since this analysis holds for every path from v_s to v_t, this investment profile outperforms investing purely on the minimum edge cut. ■

VI. Benefits of a behavioral player in multiple-defender games

We now return to the setting with multiple defenders, and consider the scenario where one of the defenders in a multi-defender network exhibits behavioral decision-making while the other defenders are non-behavioral. Through the following example, we show that the non-behavioral defenders can, in fact, *benefit* from the behavioral player's suboptimal decision-making.

Example 2: We consider the attack graph of Figs. 3 and 4. There are two players, D_1 and D_2. Player D_1 wishes to protect node 5, and player D_2 wishes to protect node 6. Note that D_1's asset (node 5) is directly on the attack path to D_2's asset (node 6). We let the total budget B_T for defending the network be $B_T = 24$, and assume that the budget distribution is asymmetric, so that player D_1 has a budget $B_1 = \frac{B_T}{5}$,

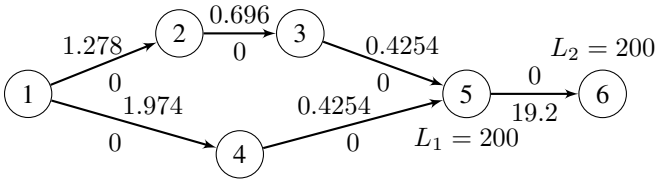

Fig. 3: Let $\alpha_1 = 1, \alpha_2 = 1$. The numbers above (below) each edge represent investments by player D_1 (D_2). Here, the non-behavioral player D_1 does not receive any investment contributions from the the other non-behavioral player D_2.

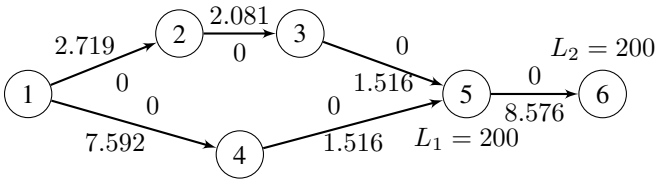

Fig. 4: Let $\alpha_1 = 1, \alpha_2 = 0.6$. The numbers above (below) each edge represent investments by player D_1 (D_2). Here, the non-behavioral player D_1 benefits from the investment contributions of the behavioral player D_2.

while player D_2 has a budget $B_2 = \frac{4B_T}{5}$. All of the optimal investments were calculated using CVX [21].

Fig. 3 shows that if both players are non-behavioral, player D_2 will not spend any money in the subnetwork of player D_1 (this could also be inferred from Theorem 2). On the other hand, Fig. 4 shows that if player D_2 is a behavioral player, she will put some investment on edges outside of her minimum edge cut (see Proposition 1), which in this example, corresponds to player D_1's edges. Therefore, player D_1's subnetwork will benefit due the behavioral decision-making of player D_2.

It is also worth considering the total expected loss of the game at equilibrium, given by $E_T = E_1 + E_2$. For this example, when both players are non-behavioral $E_T = 18.14$, while $E_T = 0.36$ if player 2 is behavioral (with $\alpha_1 = 1, \alpha_2 = 0.6$). This considerable drop in the total expected loss shows that the behavioral player's contributions to the non-behavioral player's subnetwork may also be beneficial to the overall welfare of the network, especially under budget asymmetries or if player D_1's asset is more valuable. □

VII. Discussion and Conclusion

This paper presents a game-theoretic framework that accounts for behavioral attitudes of defenders in the security of cyber-physical systems. Specifically, we proved the existence of a PNE for such Behavioral Security Games between multiple defenders. We also showed how nonlinear perceptions of attack probabilities affect the security investments made by defenders to protect their assets. In particular, non-behavioral players find it optimal to invest only on a minimum edge cut, whereas behavioral players do not. In the case of multiple defenders, we illustrated that a non-behavioral player can benefit from the presence of a behavioral player, as the latter may make (sub-optimal) investments that lead to increased protection of certain edges in the former's network. In our defense resource allocation game, the most relevant behavioral decision aspect is the probability weighting of outcomes. However, there are also various behavioral characteristics that can affect the perceived values of gains and losses [10], which we leave for future work. Moreover, it would be interesting to explore settings with strategic attackers.

References

[1] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security - A survey," *CoRR*, vol. abs/1701.04525, 2017. [Online]. Available: http://arxiv.org/abs/1701.04525

[2] V. Shandilya and S. Shiva, "On a generic security game model," *International Journal of Communications, Network and System Sciences*, vol. 10, no. 07, p. 142, 2017.

[3] A. Laszka, M. Felegyhazi, and L. Buttyan, "A survey of interdependent information security games," *ACM Computing Surveys (CSUR)*, vol. 47, no. 2, p. 23, 2015.

[4] T. Alpcan and T. Başar, *Network security: A decision and game-theoretic approach*. Cambridge University Press, 2010.

[5] A. Sanjab and W. Saad, "Data injection attacks on smart grids with multiple adversaries: A game-theoretic perspective," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2038–2049, 2016.

[6] J. Homer, S. Zhang, X. Ou, D. Schmidt, Y. Du, S. R. Rajagopalan, and A. Singhal, "Aggregating vulnerability metrics in enterprise networks using attack graphs," *Journal of Computer Security*, vol. 21, no. 4, pp. 561–597, 2013.

[7] P. Naghizadeh and M. Liu, "Opting out of incentive mechanisms: A study of security as a non-excludable public good," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 12, pp. 2790–2803, 2016.

[8] R. J. La, "Interdependent security with strategic agents and cascades of infection," *IEEE/ACM Transactions on Networking (TON)*, vol. 24, no. 3, pp. 1378–1391, 2016.

[9] A. R. Hota, A. A. Clements, S. Bagchi, and S. Sundaram, "A game-theoretic framework for securing interdependent assets in networks," in *Game Theory for Security and Risk Management*. Springer, 2018, pp. 157–184.

[10] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica: Journal of the econometric society*, pp. 263–291, 1979.

[11] A. R. Hota and S. Sundaram, "Interdependent security games on networks under behavioral probability weighting," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 262–273, March 2018.

[12] A. R. Hota, S. Garg, and S. Sundaram, "Fragility of the commons under prospect-theoretic risk attitudes," *Games and Economic Behavior*, vol. 98, pp. 135–164, 2016.

[13] A. Sanjab, W. Saad, and T. Başar, "Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game," in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–6.

[14] D. Prelec, "The probability weighting function," *Econometrica*, pp. 497–527, 1998.

[15] Y. Baryshnikov, "IT security investment and Gordon-Loeb's 1/e rule." in *Workshop on Economics and Information Security (WEIS)*, 2012.

[16] L. A. Gordon and M. P. Loeb, "The economics of information security investment," *ACM Transactions on Information and System Security (TISSEC)*, vol. 5, no. 4, pp. 438–457, 2002.

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[18] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," *Recent advances in learning and control*, pp. 95–110, 2008.

[19] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica: Journal of the Econometric Society*, pp. 520–534, 1965.

[20] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Upper Saddle River, 2001, vol. 2.

[21] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.